DOCUMENT RESUME

ED 093 992 TM 003 848

AUTHOR Jacko, Edward J.; Huck, Schuyler W.
TITLE The Effect of Varying the Response Format on the Statistical Characteristics of the Alpert-Haber Achievement Anxiety Test.
PUB DATE Apr 74
NOTE 15p.; Paper presented at the Annual Meeting of the American Educational Research Association (59th, Chicago, Illinois, April 1974)

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS Achievement Tests; *Anxiety; College Students; Comparative Analysis; *Response Mode; Test Construction; *Testing Problems; Test Reliability; Test Validity
IDENTIFIERS *Alpert Haber Achievement Anxiety Test

ABSTRACT

The Alpert-Haber Achievement Anxiety Test was developed to measure the extent to which individuals experience test anxiety. In at least two published studies, the authors claim to have used the test when in fact the response format was changed from that used in the original instrument and the "buffer" items were omitted. To investigate the possibility that these seemingly minor alterations might change the statistical characteristics of the Achievement Anxiety Test, 104 Ss were randomly assigned to each of the three versions of the Achievement Anxiety Test--the original and the two "revisions." An analysis of covariance revealed significant differences between the three forms on all three possible scales: Facilitative, Debilitative, and Facilitative-minus-Debilitative. The three versions of the Achievement Anxiety Test also differed in terms of intro-form correlation between the Facilitative and Debilitative scales and internal consistency reliability estimates. Also of interest is the finding that, within each of the three forms, the negative correlation between Facilitative and Debilitative scales caused the Facilitative-minus-Debilitative scale to yield a higher reliability coefficient than either the Facilitative scale or the Debilitative scale. (Author)

ED 093992

TM 003 848

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE EFFECT OF VARYING THE RESPONSE FORMAT ON THE STATISTICAL CHARACTERISTICS OF THE ALPERT-HABER ACHIEVEMENT ANXIETY TEST

Abstract

The Alpert-Haber Achievement Anxiety Test was developed to measure the extent to which individuals experience test anxiety. In at least two published studies, the authors claim to have used the test when in fact the response format was changed from that used in the original instrument and the "buffer" items were omitted. To investigate the possibility that these seemingly minor alterations might change the statistical characteristics of the Achievement Anxiety Test, 104 Ss were randomly assigned to each of the three versions of the Achievement Anxiety Test--the original and the two "revisions." An analysis of covariance revealed significant differences between the three forms on all three possible scales: Facilitative, Debilitative, and Facilitative-minus-Debilitative. The three versions of the Achievement Anxiety Test also differed in terms of intra-form correlation between the Facilitative and Debilitative scales and internal consistency reliability estimates. Also of interest is the finding that, within each of the three forms, the negative correlation between Facilitative and Debilitative scales caused the Facilitative-minus-Debilitative scale to yield a higher reliability coefficient than either the Facilitative scale or the Debilitative scale.

ED 093992

TM 003 848

THE EFFECT OF VARYING THE RESPONSE FORMAT ON THE STATISTICAL CHARACTERISTICS OF THE ALPERT-HABER ACHIEVEMENT ANXIETY TEST

Edward J. Jacko and Schuyler W. Huck

The University of Tennessee

In conducting original research investigations, an investigator must ask himself whether he has correctly and thoroughly described both the independent and dependent variables. This is imperative if others are to attempt replication of the study or utilization of the measuring devices. Also, when a researcher states that a particular study was repeated or a specific measuring instrument was used, no alterations should be made without a thorough detailing of these changes. Although most writers would agree that this goes without saying, violations are quite evident throughout the published literature.

The Alpert-Haber Achievement Anxiety Test (Alpert & Haber, 1960) is a case in point. The Achievement Anxiety Test was developed to indicate the extent to which an individual experiences anxiety in academic achievement test situations and whether the anxiety has a facilitative or debilitative effect on test performance. The instrument contained 28 items: nine constituted the "facilitative" scale, 10 the "debilitative" scale, and the remaining nine items were "buffer items." Each item was made up of a statement and five possible responses listed vertically, as in multiple-choice format. For example, the 16th item on the original Achievement Anxiety Test appeared as follows:

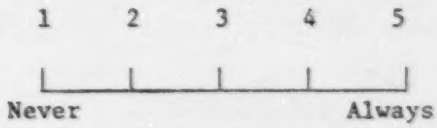
I look forward to exams.
(a) Never
(b) Hardly ever
(c) Sometimes
(d) Usually
(e) Always

Throughout the original Achievement Anxiety Test, the responses were tailored to

each statement and thus the phraseology of the responses varied across the 28-item inventory. The person taking the scale was asked to respond to each statement by selecting one of the five responses which most described his feelings.

A study by Smouse and Munz (1969) indicated that individuals stereotyped on the basis of the Achievement Anxiety Test perform differently on academic achievement tests depending on item difficulty sequencing. However, the basic Achievement Anxiety Test, as devised by Alpert and Haber, was not used. Instead, only 19 of the items constituted the modified version, the buffer items being eliminated. More importantly, the format for responding involved a five-point horizontal continuum rather than the "multiple-choice" format. Although the end points of the continuum were verbally defined with the same words or phrases that were used for the first and last responses from the original Achievement Anxiety Test, the intermediate three points on the continuum were left undefined. Thus, item 16 from the original Achievement Anxiety Test appeared in the Smouse and Munz version as follows:

I look forward to exams.



These changes to the Achievement Anxiety Test were not mentioned as such in the published articles.

Walsh used still another variation of the Achievement Anxiety Test in three recent studies. His investigations involved the relationship between test anxiety and (a) extroversion, neuroticism, and optimism (Walsh, 1968a), (b) classroom test performance (Walsh, Engbretson, & O'Brien, 1968b), and (c) psychological needs, as measured by the Edwards Personal Preference Schedule (Walsh, 1969). His modified Achievement Anxiety Test utilized a Likert-type response format for each statement with the phraseology of the five response options being constant across all items. For example, item 16 from the original Achievement Anxiety Test appeared in the

Walsh version as follows:

I look forward to exams. R ___ S ___ F ___ G ___ A ___

The directions explained that R = rarely, S = sometimes, F = frequently, G = generally, and A = almost always. On this version of the Achievement Anxiety Test, the buffer items again were omitted. As was the case with Smouse and Munz, Walsh did not indicate these changes to the original Achievement Anxiety Test.

Several additional investigators (Bauer, 1971; Hermans, Laak, & Maes, 1972; Laxter, Quarter, Kooman, & Walker, 1969; Laxter & Walker, 1970; Wittmaier, 1972) have conducted studies in which it is stated that the Alpert-Haber Achievement Anxiety Test was used. In light of the above, merely mentioning the use of the Achievement Anxiety Test leaves us in doubt as to whether the basic instrument is actually being utilized. Of course, if the different versions of the Achievement Anxiety Test produce similar results, then we need not be concerned.

With the thought that variations in the response format might cause examinees to respond differentially to the Achievement Anxiety Test, the present study was undertaken to compare the three forms of the Achievement Anxiety Test described above. More specifically, comparisons were made in terms of mean score and reliabilities for the three possible scores that could be computed from the Achievement Anxiety Test: Facilitative, Debilitative, and Facilitative-minus-Debilitative. In addition, the three forms were compared in terms of intra-form correlation between the facilitative and debilitative scores. Any significant differences would obviously negate generalizing that all forms of the Achievement Anxiety Test would equally measure test anxiety and would provide substantiation for the requirement of correct and thorough description of all variables in published studies.

METHOD

Subjects

The subjects (Ss) for this experimental study consisted of all undergraduate

students enrolled in the nine sections of an adolescent psychology course offered at The University of Tennessee during the fall academic term, 1972. A total of 312 Ss were enrolled in these classes, with an average class size of approximately 35 students.

Procedure

Prior to the day of the experiment, 312 packets of material were assembled by the investigators. The first item in each packet was a single sheet of paper requesting the student to rate himself in terms of how his test performance was affected by test anxiety. The instructions asked the student to indicate whether test anxiety prevented him from doing well on tests or helped him to achieve high exam scores, and a -5 to +5 scale was provided for the self rating. This scale was carefully defined such that -5 would indicate the maximum amount of negative test anxiety, a 0 would indicate no test anxiety at all, and a +5 would indicate the maximum amount of positive test anxiety. The "self rating" sheet was identical for all 312 packets.

The second item in each packet was an envelope containing one of the three forms of the Achievement Anxiety Test. Except for necessary modifications in the instructions to achieve uniformity, these three forms of the Achievement Anxiety Test were reproduced exactly as they had been used in previous research studies. This meant that the three forms were identical in terms of the wording of the statements to which the Ss responded, and dissimilar in terms of the response format and the inclusion of the buffer items on the original 1960 form.

The actual experiment was conducted on the first day of the academic term. In each participating class, the regular instructor associated with that particular section of the adolescent psychology course distributed the packets to the students. Since the packets had previously been arranged in random order, one-third of each class was randomly assigned to a different form of the Achievement Anxiety Test. After each student had received a packet, the instructor emphasized the importance

of completing the "self rating" before opening the envelope. In addition, the students were encouraged to be honest in completing both items in their packet since no names were requested on the single sheet of paper (i.e., the "self rating") or the envelope or the material inside the envelope.

Statistical Treatment of Data

For each of the Ss, three criterion measures were derived from the S's responses to the Achievement Anxiety Test. These were the Facilitative, the Debilitative, and the Facilitative-minus-Debilitative scores. A separate one-way analysis of covariance was used to compare the three forms of the Achievement Anxiety Test with respect to each criterion measure, with scores from the "self rating" serving as the covariate in each analysis. Since three treatment groups were involved in these covariance analyses, Tukey's post hoc multiple comparison test was used to compare the adjusted criterion means following an over-all significant F-ratio (Winer, 1971, p. 772).

Prior to using the analysis of covariance, the assumption of homogeneous within-group regression slopes was tested. In accordance with Kirk's (1968) recommendation, this assumption was tested at the .25 level of significance to protect against the possibility of a Type II Error. The assumptions of normality and homogeneity of variance were not tested since the analysis of covariance, like the analysis of variance, is robust to violations of these assumptions under the condition of equal numbers of Ss in the treatment groups.

The reliabilities for the Facilitative and Debilitative scales of each form of the Achievement Anxiety Test were estimated through an extension of the Hoyt analysis of variance reliability procedure (Hoyt & Stunkard, 1952). The reliability of the Facilitative-minus-Debilitative scores for each form of the instrument was also estimated through the use of Horst's (1966) formula for estimating the reliability of difference scores.

Finally, the correlation between the Facilitative and Debilitative scores within each form of the Achievement Anxiety Test was obtained. These three correlations were then compared with an appropriate chi square test for determining whether significant differences are present among k correlation coefficients, where $k > 2$.

All major analyses of data were conducted at The University of Tennessee Computer Center through the use of MANOVA and SAS library programs. For all inferential statistical tests (except those dealing with the assumption of homogeneous regression slopes in the covariance analyses), the present investigators made an a priori decision to use, and to report, the .05 level of significance.

RESULTS

Means and standard deviations on the "self rating" (the covariate) and the Facilitative, Debili ative, and Faciliative-minus-Debilitative scales of the Achievement Anxiety Test for each of the three treatment groups are presented in the first three columns of Table 1. Also presented in these columns are the adjusted means for each treatment groups on each scale.

Insert Table 1 about here

Prior to using the analysis of covariance to compare the three treatment groups on each scale of the Achievement Anxiety Test, a test was made to determine whether the assumption of homogeneous within-group regression slopes was tenable. The results of these tests are presented in the fourth column of Table 1, and in each case the assumption was not rejected ($p > .25$). This outcome permitted the present investigators to use the conventional covariance analysis in comparing the adjusted means.

As indicated in the last column of Table 1, significant differences ($p < .05$) were found among the three treatment groups on all three scales of the Achievement

Anxiety Test. With respect to the Facilitative scale, Tukey's multiple comparison test revealed significant differences ($p < .05$) between all pairwise comparisons. On the Debilitative and Facilitative-minus-Debilitative scales, Tukey's test indicated significant differences between Forms #1 and #2 and between Forms #2 and #3 but not between Forms #1 and #3.

The reliability estimates for the Facilitative, Debilitative, and Facilitative-minus-Debilitative scales on each of the three forms of the Achievement Anxiety Test are presented in Table 2. The reliability coefficients ranged from .65 to .76 on the Facilitative scale, from .76 to .86 on the Debilitative scale, and from .83 to .88 on the Facilitative-minus-Debilitative scale. With respect to each of the three forms of the Achievement Anxiety Test, the reliability of the Facilitative-minus-Debilitative scale was higher than the reliability for either the Facilitative scale or the Debilitative scale. This increase in the reliability of the difference scores is attributable to a negative correlation between the Facilitative scores and the Debilitative scores: -.66 for Form #1, -.42 for Form #2, and -.60 for Form #3. When compared statistically, significant differences were shown to exist among these three correlation coefficients ($\chi^2 = 6.001$, $df = 2$, $p < .05$).

Insert Table 2 about here

DISCUSSION

The results of this research investigation demonstrate that sometimes even seemingly minor variations of the original Alpert-Haber Achievement Anxiety Test can yield scores having different means, reliability estimates, and intra-form correlations between the Facilitative and Debilitative scores. In addition to the finding of differences among forms, this study reveals that the negative correlation between the Facilitative and Debilitative scales within any one form of the Achieve-

ment Anxiety Test causes the reliability of the Facilitative-minus-Debilitative scale to be higher than the reliability of the Facilitative or the Debilitative scale. This finding suggests that the Facilitative-minus-Debilitative scale is the best over-all measure of test anxiety instead of the Facilitative or Debilitative scale as some authors have implied.

Two comments concerning the research design used in this study are worth noting. First, the three forms of the Achievement Anxiety Test varied along two dimensions, response format and inclusion vs. omission of the buffer items. The potential confounding between these two variables could have been avoided through the use of a 2x3 factorial design. However, such a design would have necessitated the construction of three new forms of the Achievement Anxiety Test. Since the only purpose of this study was to compare the three forms of the Achievement Anxiety Test that have been used in published research investigations, a one-factor design was employed. Although the authors were willing to find differences without knowing the specific cause, the observed differences between the two revised versions of the Achievement Anxiety Test, neither of which had buffer items, clearly indicate that the factor of response format, by itself, has an effect on the instrument.

The second comment related to the design concerns the use of the self rating as the covariate. The present authors recognize that there might be differences between the way individuals respond to the Achievement Anxiety Test depending upon whether or not they are first asked to rate themselves on test anxiety. Thus, the results obtained in this study with respect to any particular form of the Achievement Anxiety Test might not be replicable with a similar group of Ss who would be asked to complete just the Achievement Anxiety Test and not the self rating. In spite of this possibility, the administration of the self rating prior to the Achievement Anxiety Test in the present study cannot be considered as a plausible rival hypothesis for the observed differences between the three forms of the Achievement

Anxiety Test. Since all Ss were exposed to the self rating first, a sensitizing effect--if there was one--was controlled.

In conclusion, the differential results obtained with three variations of the same instrument in this investigation do not indicate that slight alterations in an instrument will always create a different score distribution. The results do, however, argue for comprehensive description of variables in a study to eliminate error in replication, and also for an indication of any changes that have been made to the measuring instruments used within the study.

References

Alpert, R., & Haber, R. N. Anxiety in academic achievement situations. Journal of Abnormal and Social Psychology, 1960, 61, 207-215.

Bauer, D. H. The effect of test instructions, test anxiety, defensiveness, and confidence in judgment on guessing behavior in multiple-choice test situations. Psychology in the Schools, 1971, 8, 208-215.

Hermans, H. J. M., Laak, J. J. F., & Maes, P. C. J. M. Achievement motivation and fear of failure in family and school. Developmental Psychology, 1972, 6, 520-528.

Horst, P. Psychological measurement and prediction. Belmont, California: Wadsworth, 1966.

Hoyt, C. J., & Stunkard, C. L. Estimation of test reliability for unrestricted item scoring methods. Educational and Psychological Measurement, 1952, 12, 756-758.

Kirk, R. E. Experimental design: Procedures for the behavioral sciences. Belmont, California: Brooks/Cole, 1968.

Laxter, R. M., Quarter, J., Kooman, A., & Walker, K. Systematic desensitization and relaxation on high-test-anxious secondary school students. Journal of Counseling Psychology, 1969, 16, 446-451.

Laxter, R. M., & Walker, K. Counter-desensitization versus relaxation in the desensitization of test anxiety. Journal of Counseling Psychology, 1970, 17, 431-436.

Smouse, A. D., & Munz, D. C. Item difficulty sequencing and response style: A follow-up analysis. Educational and Psychological Measurement, 1969, 29, 469-472.

Walsh, R. P. Test-taking anxiety and psychological needs. Psychological Reports, 1969, 25, 83-86.

Walsh, R. P. Some correlates of test-taking anxiety. Psychological Reports, 1968, 22, 449-450.

Walsh, R. P., Engbretson, R. O., & O'Brien, B. A. Anxiety and test-taking behavior. Journal of Counseling Psychology, 1968, 15, 572-575.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1971.

Wittmaier, B. C. Test anxiety and study habits. Journal of Educational Research, 1972, 65, 352-354.

Table 1

Comparison of the Three Forms of the Achievement Anxiety Test (AAT) Using the Analysis of Covariance

	TREATMENT GROUPS			Statistical Comparison of the 3 Forms	
	Form #1 (n = 104)	Form #2 (n = 104)	Form #3 (n = 104)		
Self Rating	$\bar{X}$ = -.990 SD = 2.799	$\bar{X}$ = -.990 SD = 2.631	$\bar{X}$ = +.394 SD = 2.736	Test of Assumption of Equal Regression Slopes	Analysis of Covariance Comparing Adjusted Means
Facilitative Scale of AAT	$\bar{X}$ = 24.933 SD = 5.079 $\bar{X}_{adj}$ = 25.063	$\bar{X}$ = 19.288 SD = 6.029 $\bar{X}_{adj}$ = 19.419	$\bar{X}$ = 23.962 SD = 5.039 $\bar{X}_{adj}$ = 23.701	F = 0.932 df = 2/306 $p > .25$	F = 40.4 df = 2/308 $p < .05$
Debilitative Scale of AAT	$\bar{X}$ = 29.356 SD = 6.395 $\bar{X}_{adj}$ = 29.202	$\bar{X}$ = 23.173 SD = 6.606 $\bar{X}_{adj}$ = 23.019	$\bar{X}$ = 29.904 SD = 6.733 $\bar{X}_{adj}$ = 30.212	F = 0.425 df = 2/306 $p > .25$	F = 46.7 df = 2/308 $p < .05$
Facil. minus Debil. Scale	$\bar{X}$ = -4.433 SD = 10.409 $\bar{X}_{adj}$ = -4.147	$\bar{X}$ = -3.982 SD = 10.627 $\bar{X}_{adj}$ = -3.696	$\bar{X}$ = -5.942 SD = 10.545 $\bar{X}_{adj}$ = -6.513	F = 0.827 df = 2/306 $p > .25$	F = 3.05 df = 2/308 $p < .05$

Form #1 = original AAT; Form #2 = Smouse & Munz version, Form #3 = Walsh version

Table 2

Reliability Estimates for the Three Scales of the Achievement Anxiety Test

	Form #1 (n = 104)	Form #2 (n = 104)	Form #3 (n = 104)
Facilitative	.73	.76	.65
Debilitative	.86	.76	.82
Facilitative minus Debilitative	.88	.83	.85